DOCUMENT RESUME

ED 322 215 TM 015 461

AUTHOR	Biester, Thomas W.
TITLE	Trends in School Improvement: Regional Test Results, 1984-1989.
INSTITUTION	Research for Better Schools, Inc., Philadelphia, Pa.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Apr 90
CONTRACT	400-86-0003
NOTE	45p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Achievement Tests; Data Analysis; Educational Improvement; *Educational Trends; Elementary Secondary Education; Mathematics Achievement; *Needs Assessment; Reading Achievement; *Regional Programs; Standardized Tests; Testing Programs; Test Norms; *Test Results; Trend Analysis
IDENTIFIERS	Pennsylvania; Scholastic Aptitude Test

ABSTRACT

As part of an overall regional needs assessment, an overview is provided of testing programs in the mid-Atlantic region, consisting of: the District of Columbia; Delaware; Maryland; New Jersey; and Pennsylvania. The testing programs vary widely by objective and approach. One of the major needs assessment activities is the examination of student performance data related to designated learning goals for each jurisdiction. From the diverse types of data, analysis has yielded individual state performance trends and regional performance trends. Scores were converted to a standard score metric to enable meaningful comparisons of gross student performance trends. Major conclusions include the following: (1) reading achievement scores fluctuated from 1984 to 1989, with improvement for elementary and secondary grades and stability or decline for intermediate grades; (2) for the whole period, there was overall improvement in mathematics, with some declines in 1988 and 1989; (3) overall results for 1989, while generally positive, were not as strong as those for previous years; (4) students in the region generally performed above national norms on standardized tests, and scores were particularly high for the elementary school level; (5) most students met minimum proficiency scores established by the states; and (6) scores on the Scholastic Aptitude Test (SAT) tended to be below the national average, but regional SAT scores improved from 1988 to 1989. Each state is in the process of reviewing testing programs to ensure that accurate indicators of achievement are available. Twelve tables provide study data. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ellen Newcombe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TRENDS IN SCHOOL IMPROVEMENT:

REGIONAL TEST RESULTS

Thomas W. Biester

Research for Better Schools
444 North Third Street
Philadelphia, Pennsylvania 19123

April 1990

This publication is based on work sponsored, wholly, or in part, by the Office of of Educational Research and Improvement (OERI), Department of Education, under Contract Number 400-86-0003. The content of this publication does not necessarily reflect the views of OERI, the Department, or any other agency of the U.S. Government.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
Overview of Testing Programs	1
Previous RBS Reports on Regional Achievement Trends	6
Analysis Approach	7
ANALYSIS OF INDIVIDUAL STATE PERFORMANCE TRENDS	10
District of Columbia	10
Delaware	12
Maryland	14
New Jersey	16
Pennsylvania	18
ANALYSIS OF REGIONAL BASIC SKILLS ACHIEVEMENT	22
Regional Reading Trends	22
Regional Mathematics Trends	28
CONCLUSIONS	32
REFERENCES	34

Introduction

This report represents part of an overall regional needs assessment effort undertaken by Research for Better Schools (RBS). Needs assessment is an integral part of the laboratory's overall planning and development process. Needs assessment activities are intended to determine the regional responsiveness of current laboratory programs, to redirect programs or establish new programs, and to provide feedback to both the Office of Educational Research and Improvement (OERI) and state departments of education regarding empirically derived needs. One of the major needs assessment activities is the examination of student performance data related to designated learning goals for each of the jurisdictions in the mid-Atlantic region (District of Columbia, Delaware, Maryland, New Jersey, and Pennsylvania).

This introduction provides an overview of the five testing programs, describes results from previous trend analysis reports, and outlines the analysis approach which was used this year. Subsequent sections of the report present performance results by jurisdiction, a summary of basic skills results across jurisdictions, and conclusions.

Overview of Testing Programs

Each education agency in the region has a mandated testing program to assess student performance, particularly in the basic skills areas. Table 1 presents a brief overview of the testing programs discussed in this report.

Table 1

Overview of Testing Programs

Jurisdiction	Testing Program	General Description	Grades Tested	Score Types/ Norming	Testing Sample	Primary Unit of Analysis/Reporting
District of Columbia	Annual testing	CTBS in reading, mathematics, language, science, social studies, and reference skills	3, 6, 8, 9, 11	Grade equivalents and percentiles/ national norms	All students, each year	District and school levels
Delaware	Delaware Educational Assessment Program (DEAP)	Stanford Achievement Test (SAT) in reading, mathematics, language, spelling, listening, study skills, science, social science, thinking skills, using information	1-8, 11	Normal curve equivalents (NCEs)/ national norms	All districts, each year	State, district, and school levels
Maryland	Maryland Accountability Testing Program: Norm-Referenced Component; Criterion-referenced Component: Project Basic, high school competency examinations - data not available for report	CAT in reading, mathematics, language	3, 5, 8	Grade equivalent scores, scale scores, and NCEs/ national norms	All districts and schools, each year	State, district (county), and school levels available/national

Table 1 (continued)

Jurisdiction	Testing Program	General Description	Grades Tested	Score Types/ Norming	Testing Sample	Primary Unit of Analysis/Reporting
New Jersey	High School Proficiency Test (HSPT); local districts select own standardized test for elementary grades -- data not available for report	State developed reading, mathematics, and writing tests	9; retests in subsequent years for those who fail	HSPT scores (percent correct equated to 1984 test sample)/statewide passing scores	All districts, each year	State, district, and school levels
Pennsylvania	Educational Quality Assessment (EQA)	State-developed test with numerous subtests; new battery developed in 1985; includes subtests in reading, writing, mathematics, analytical thinking, social studies, arts and humanities, science and technology, environment, self-concept, health practices, and health knowledge	1985: 5,8,11; 1986-88: 4, 6, 7, 9, 11; not administered in 1989	Raw scores and percentiles/statewide norms, vary year to year	Voluntary; different sample from year to year	School level
	Testing for Essential Learning and Literacy Skills (TELLS)	State-developed testing and remedial instruction program begun in 1984; "minimum" basic skills exams in reading and mathematics; items chosen from nationally standardized item pool each year (Chas. Merrill Publishing)	3, 5, 8	Raw scores, percent correct; percentage above/below state-established cutoff scores/estimated national norms	Each school, each year	State, district, and school levels

As Table 1 indicates, the testing programs vary widely in objective and approach. Several programs use commercially available standardized tests to compare student achievement with national norms (DC, DE, and MD). Others are locally constructed (NJ, PA) and are used more like criterion-referenced tests. The latter focus on a wide range of educational goals, assess minimum basic skills or higher level proficiencies, and use local norms or standards, although new testing packages in both states incorporate some elements of a nationally based norm-referenced system. For example, local districts in New Jersey now select a commercially available standardized test of their own choice for assessment of elementary grade students.

District of Columbia. The DC Public Schools annually administer six subtests of the Comprehensive Tests of Basic Skills (CTBS) to students in five grades. Student progress in terms of national norms (grade equivalents) are reported on a district and school basis.

Delaware. Delaware switched standardized achievement tests from the CTBS to the Stanford Achievement Test (SAT) in 1989. The battery included eight to ten subtests, administered to all elementary and middle grades (1-8), as well as one secondary grade (11). Results are reported in average normal curve equivalent (NCE) scores at the state, district, and school levels. Prior to the SAT, the state used the California Achievement Tests (1978-1983) and the CTBS (1984-88).

Maryland. The norm-referenced component of the Maryland Accountability Testing Program uses three subtests of the CAT for three grades (3, 5, and 8). Like the District of Columbia, average grade equivalent scores are reported on district (county) and school levels. Maryland also administers a series of

criterion-referenced functional examinations to high school students (Project Basic), but results were not available for this report.

New Jersey. The New Jersey testing program has changed considerably over the past few years. Previously, the Minimum Basic Skills Testing Program (MBS) was developed to detect deficiencies in minimum basic skills. However, since deficiencies in higher level skills could not be detected and most students could correctly answer most items by 1983, the state decided to move towards a system which utilized national standardized achievement tests at the elementary grade levels (each district choosing its own test battery), and a new higher level achievement test which secondary students must pass to graduate from high school, the New Jersey High School Proficiency Test (HSPT). The HSPT is administered to all ninth grade students, and those who fail retake the examination in subsequent years. The HSPT requirement became effective in 1986 for all incoming ninth graders. Baseline data were collected in 1984 and 1985, with subsequent scores equated to the 1984 test sample.

Pennsylvania. The Pennsylvania testing program also has been in a period of transition for the past few years. The Educational Quality Assessment (EQA) Program was designed to address the state's goals for quality education. As such, it represents a comprehensive test battery with a wide variety of subtest areas. School district participation in the EQA is voluntary. An essentially new EQA test battery was developed in 1985 to replace the 1978-1984 battery. The state decided not to administer the EQA program during the 1988-1989 school year and so no results are reported this year.

As part of the state's educational reform package announced in 1983, the Testing for Essential Learning and Literacy Skills (TELLS) examination program was designed as an "early warning system" aimed at diagnosis and remediation of basic reading and mathematics problems early in students' school careers. Since the TELLS examination is now mandated for grades 3, 5 and 8, the EQA was offered to grades 4, 6, 7, 9, and 11 instead of grades 5, 8, and 11, as in years prior to 1986.

Previous RBS Reports on Regional Achievement Trends

Statewide achievement trends have been described in numerous RBS reports presented in prior years (Biester & Buttram, 1986; Biester & Dusewicz, 1982, 1983; Dusewicz & Kenney, 1984). Results for these reports were presented with 1978 test scores as a baseline in most states. In addition, regional trends were compared to findings from the National Assessment of Educational Progress (NAEP). Reading trends were generally positive at all grade levels during the period from 1978-1984. These results were similar to NAEP findings for this time period. Students at all grade levels throughout the region also showed considerable improvement in mathematics from 1978-1984, unlike NAEP findings which showed little change during this period. Overall, students across the region performed much higher than national norms on standardized scores. Average scores for early grades tended to be higher than secondary grade scores. Performance for many urban areas was relatively poor, although there were some exceptions. Students across the region demonstrated a high degree of mastery of minimum basic skills.

The 1987 and 1988 reports on achievement trends (Biester, Herzel-Colvson, & Buttram, 1987; Biester & Rioux, 1988) used 1984 test scores as a baseline

because of changes in state testing programs since 1984. Trends from 1984-1988 generally followed those exhibited from 1978-1984. Reading trends were fairly stable while there was some improvement in mathematics during the period. Overall, students across the region performed much higher than national norms on standardized tests and demonstrated a high degree of mastery of minimum basic skills. Analyses indicated that scores and score trends were probably inflated because of factors related to old test norms and increasing familiarity with test content, similar to findings noted by Cannell (1988), Shepard (1989), and others.

Analysis Approach

Like previous reports, the RBS assessment of student performance in 1989 consisted of a secondary analysis of existing data available from the five testing programs. As in prior years, the major focus of the report was an analysis of year-to-year trends in students' basic skills achievement (reading and mathematics) both within jurisdictions and across the region. The analysis approach for each type of comparison is described below.

Individual state performance trends. Trends at the elementary, intermediate, and secondary school levels were analyzed. Results from grades 3, 5, 8, and 11 were used where scores were available. In some cases, no data were available at a particular grade level (e.g., New Jersey elementary and intermediate levels, Maryland secondary level).

Since major changes in the testing programs for several states have occurred in the past few years, 1984 performance was chosen as a baseline year for the 1989 analyses. Average 1984 scores were available for the District of Columbia, Maryland, and New Jersey. For Pennsylvania, 1985-1989 TELLS data

were appropriate for analysis; no analyses were performed related to EQA because the state decided to not administer the EQA battery in 1989. No baseline data were appropriate for Delaware because of the change in test form in 1989. Average score changes from the baseline year were examined to determine if performance was stable, improving, or declining. Various score types were analyzed, including raw scores, percent correct scores, and national normal curve equivalent (NCE) scores, depending upon their availability for the specific testing program. In addition, converted standard scores as described below, were analyzed within each state.

Regional performance trends. Although the analysis of trends within a state is relatively straightforward, the determination of regional trends is somewhat difficult. This difficulty arises because different tests are used, different grades are involved, and different score types and norms are used in the various state testing programs. An analysis approach was developed in prior reports to address these problems in order to make meaningful regional assessments, and to enable more meaningful year-to-year comparisons within states where there were no national norms.

This analysis approach, which was used again this year, requires the conversion of existing test scores to a common testing metric. Baseline scores, (e.g., mean test scores for 1984 or the first available year) were arbitrarily set as a standard score of 50. Mean scores for subsequent years were converted to standard scores (with a mean of 50 and a standard deviation of 21.06) using the baseline score as a reference point. This resulted in an equal interval scale with a hypothetical range from 1 to 99 (assuming that statewide averages would not change by more than two and one-half standard deviations). Comparisons of scores from year to year thus represented changes

in student performance from that baseline year in terms of standard deviation units. However, direct comparisons cannot be made across states because each test differs with regard to context, difficulty levels, norming samples, and other psychometric properties.

To summarize, the purpose of the conversion of scores to the standard score metric is to enable meaningful indications of gross student performance trends only. These gross trends show if student performance in particular achievement areas is stable, improving, or declining. For example, are the 1989 reading scores for third grade students the same, better, or worse than scores for 1984 third grade students? The analyses cannot show whether students in one state are performing better or worse than students in another state.

Analysis of Individual State Performance Trends

This section describes student performance trends separately for each jurisdiction in the region. The data tables summarize the major findings for each of the content areas and grade levels addressed in this study. More detailed findings are available in reports and tables for each of the specific testing programs from each of the five jurisdictions. Performance data in terms of the standard scores calculated for this report and appropriate normative and/or raw score data are presented.

District of Columbia

NCE and converted standard scores for the 1984-1989 administrations of the Comprehensive Test of Basic Skills (CTBS) in the District of Columbia are presented in Table 2. Form S, an early version of the CTBS, was administered from 1984 through 1986 and Form U, a recently published version, was administered from 1987 through 1989. The 1984 through 1986 scores shown in Table 2 were converted from Form S mean scores to equivalent Form U mean scores in order to meaningfully interpret trends. Form S norms were based on a sample of students from 1973, whereas Form U norms are somewhat more current, based on a sample from 1981. Previous RBS achievement trend reports noted the phenomenon of apparent declines in performance when more current testing forms are used. Findings have been similar to those noted by Cannell (1988).

Table 2 presents districtwide results for reading, mathematics, and language. The effects of the change in test form are minimized in Table 2 since 1984 through 1987 scores (Form S) have been converted to Form U scores using conversion tables provided by the publisher.

Table 2

District of Columbia Achievement Trends
Comprehensive Test of Basic Skills
NCEs and (Converted Standard Scores)

Subtest	1984	1985	1986	1987	1988	1989
Grade 3						
Reading	47 (50)	45 (48)	48 (51)	46 (49)	48 (51)	50 (53)
Mathematics	52 (50)	50 (48)	51 (49)	51 (49)	54 (52)	58 (56)
Language	48 (50)	46 (48)	48 (50)	48 (50)	52 (54)	54 (56)
Grade 6						
Reading	48 (50)	48 (50)	48 (50)	47 (49)	48 (50)	50 (52)
Mathematics	54 (50)	54 (50)	53 (49)	55 (51)	56 (52)	59 (55)
Language	49 (50)	50 (51)	49 (50)	50 (51)	51 (52)	53 (54)
Grade 8						
Reading	39 (50)	39 (50)	43 (54)	40 (51)	43 (54)	43 (54)
Mathematics	42 (50)	46 (54)	47 (55)	46 (54)	47 (55)	48 (56)
Language	43 (50)	44 (51)	44 (51)	43 (50)	46 (53)	47 (54)
Grade 11						
Reading	29 (50)	35 (56)	38 (59)	40 (61)	39 (60)	43 (64)
Mathematics	38 (50)	39 (51)	40 (52)	43 (55)	43 (55)	44 (56)
Language	35 (50)	39 (54)	40 (55)	42 (57)	43 (58)	45 (60)

Note: Form S scores from 1984 through 1986 were converted to Form U score equivalents (CTBS norms tables).

At the elementary grade level, third grade scores on all subtests were fairly stable from 1984 to 1987, but improved substantially in 1988 and again in 1989. One might consider 1987 as a more reasonable baseline since that was the year that the new tests were first administered. The average NCE score in reading was at the national average in 1989, an improvement of four NCE points since 1987. Scores in mathematics and language were higher than national norms, again with positive improvement trends since 1987.

At the intermediate grade level, trends for sixth grade students paralleled those for third grade students. Average scores in all subtests improved from 1988 to 1989. Sixth grade reading scores are at the national norm, language scores are slightly above, and mathematics scores are considerably above the national average. Trends for eighth grade students are different, however. Scores on all subtests remain below national averages, despite a modest increases from 1987 to 1988. Reading scores remain very low, with an average NCE of 43, identical to last year. Language and mathematics scores showed modest gains in 1989 and are gradually approaching the national norm.

Scores at the secondary level for grade 11 students increased in 1989, but remain considerably below the national norm. Reading scores improved four NCE points in 1989 and have increased substantially since 1984, when average scores were extremely low. Mathematics and language scores improved slightly in 1989 and, again, gains were substantial since very low scores were shown in 1984.

Delaware

Delaware used the Comprehensive Test of Basic Skills (CTBS) as part of the Delaware Educational Assessment Program from 1984 through 1988, using 1981 national norms provided for the CTBS. In 1989, the state decided to switch to the Stanford Achievement Test (SAT), Eighth Edition. The SAT was normed in 1988, so more current, meaningful comparisons could be made between Delaware students and those nationwide. Various support services were provided to local districts and parents to ease the transition to the new test. These included videotapes, mini-grants, publications, and workshops for orientation

to the test's content and structure. The SAT contains subtests for listening, reading, mathematics, language, spelling, study skills, science, social science, and higher order thinking skills, including using information and thinking skills. Not all subtests are available at each grade level. Different test levels are provided for each grade.

As the state's Summary Report for Statewide Test Results notes, it is extremely difficult to make meaningful comparisons with test results from prior years. Average CTBS scores from 1984 through 1988 were relatively stable at all grade levels, with slight improvement each year. Average NCEs for all grade levels in 1988 were much higher than national norms (from 1981), with scores for all subtests in the range from approximately 58 to 65.

Average NCEs from the 1989 SAT are presented in Table 3 for grades 3, 5, 8, and 11. Scores on each subtest at each grade level are above the national norms. However, 1989 scores are lower than those reported for the CTBS in 1988. This is expected and doesn't imply that performance has declined, rather, it is almost certainly an artifact of the different, more recent test norms. Averages are very consistent across subtests, with most elementary level scores around 55 and intermediate and secondary scores slightly lower at about 53. Overall, Delaware students performed very well in relation to national norms, with reasonably high scores in basic skills areas as well as in higher order thinking skills. It will be interesting to examine trends next year, after teachers and students have become more familiar with test structure and content.

Table 3

1989 Delaware Achievement Results
Stanford Achievement Test (SAT)
Average NCE Scores

Subtest	Grade 3	Grade 5	Grade 8	Grade 11
Reading	55	54	54	53
Mathematics	54	53	53	53
Language	57	55	54	53
Spelling	55	54	53	52
Study Skills	58	55	53	53
Listening	52	52	50	53
Using Information	57	54	53	53
Thinking Skills	55	53	53	53
Science	-	-	-	52
Social Science	-	-	-	53

Note: Results are for regular education students; results for special education students are excluded.

Maryland

NCEs and converted standard scores for the California Achievement Test (CAT) given in Maryland from 1984-1989 are shown in Table 4 for grades 3, 5, and 8. Maryland students are tested in the fall, so that data presented as the 1984 school year data are actually the result of fall 1983 testing.

At the elementary level, grade 3 reading scores remained stable from 1984 through 1986, with slight improvement in 1987. The average 1989 reading score improved from the 1987 and 1988 levels. Grade 5 reading scores dropped sharply from 1984 to 1985, improving slightly in 1986 and 1987, and remaining relatively stable in 1988 and 1989. Mathematics scores for both grades 3 and 5 remained stable from 1984 through 1986, and improved moderately in 1987 and 1988. The 1989 scores were similar to last year's scores.

Table 4

Maryland Achievement Trends
California Achievement Test
NCEs and (Converted Standard Scores)

Subtest	1984	1985	1986	1987	1988	1989
Grade 3						
Reading	55 (50)	55 (50)	55 (50)	57 (52)	57 (52)	59 (54)
Math	59 (50)	59 (50)	59 (50)	64 (55)	65 (56)	64 (55)
Grade 5						
Reading	64 (50)	57 (43)	58 (44)	60 (46)	59 (45)	59 (45)
Math	60 (50)	60 (50)	60 (50)	63 (53)	64 (54)	64 (54)
Grade 8						
Reading	60 (50)	58 (48)	58 (48)	60 (50)	60 (50)	60 (50)
Math	58 (50)	57 (49)	59 (51)	61 (53)	62 (54)	62 (54)

At the intermediate level, grade 8 reading and mathematics scores dropped slightly in 1985, but improved by 1987. Reading scores returned to the 1984 level in 1987 and remained constant in 1988 and 1989. Mathematics scores improved from 1985 through 1988, and remained stable in 1989.

In summary, Maryland achievement test scores have remained fairly stable between 1984 and 1986 (except for grade 5 reading), with slight (eighth grade reading) to relatively large (third grade mathematics) improvements in 1987. The 1988 and 1989 mathematics and reading scores remained relatively stable when compared with previous years, with most scores fluctuating by only one point. Grade 3 reading scores were slightly higher in 1989. Despite steady improvement, however, grade 5 reading scores still have not regained their 1984 level after dropping sharply in 1985.

Basic skills achievement for Maryland students continues to be well above national norms in reading and mathematics at all grade levels. Average NCE scores were approximately 60 in reading and 64 in mathematics. however, it must be noted that these scores are based on norms generated in 1977. A test with more recent norms is currently being considered.

New Jersey

From 1978 through 1982, the New Jersey Minimum Basic Skills Test (MBS) was administered to students in grades 3, 6, 9, and 11. Performance trends showed sharp improvements in reading and mathematics across the years. By 1982, most students were correctly answering most test items (e.g., 90-93 percent of reading items and 85-86 percent of mathematics items). Student performance was impressive. After 1982, the test was used as a graduation proficiency examination for ninth grade students and reexaminees who had previously failed. Results of the MBS testing program have been documented in previous RBS achievement trend reports.

The New Jersey High School Proficiency Test (HSPT) was first administered to ninth grade students in 1984. The HSPT replaced the MBS as a graduation requirement for ninth graders in 1986. Results of the HSPT from 1984 through 1989 for ninth grade students are presented in Tables 5 and 6. As Table 5 indicates, achievement levels in reading, mathematics, and writing have increased steadily with each test administration. The HSPT raw scores show that students correctly answered a very high percentage of items, particularly in reading (94 percent) and writing (87 percent). The HSPT trends parallel those for the MBS tests from 1978-1982. While one would expect average scores to reach a ceiling level in subsequent years since they are currently so high,

Table 5

New Jersey Achievement Trends - 9th Grade Students
High School Proficiency Test
Percent Correct and (Converted Standard Scores)

Subtest	1984	1985	1986	1987	1988	1989
Reading	77.3	79.9	85.0	88.3	90.6	94.3
	(50)	(53)	(59)	(63)	(66)	(70)
Mathematics	62.6	64.4	71.6	74.3	77.0	80.5
	(50)	(52)	(59)	(62)	(64)	(68)
Writing	-	80.0	82.2	84.7	85.9	87.4
		(50)	(53)	(57)	(59)	(61)

Table 6

New Jersey Achievement Trends - 9th Grade Students
High School Proficiency Test
Percent Passing HSPT

Subtest	1986	1987	1988	1989	Change 1988-89	Change 1986-89
Reading	82.6	91.0	93.4	97.8	+4.4	+15.2
Mathematics	71.8	77.0	81.8	87.9	+6.1	+16.1
Writing	76.5	85.9	90.7	93.7	+3.0	+17.2

scores continued to rise substantially, even in 1989. Average reading scores seem to be approaching a perfect score. As reported in Table 6, a very high percentage of ninth grade students passed the examinations on their first attempt. Pass rates for 1989 were 98 percent, 88 percent, and 94 percent for reading, mathematics, and writing, respectively. The percentage of students passing each subtest improved each successive year, and gains continued despite the anticipated ceiling. The percentage of students passing all three sections on their first attempt improved from 71 percent in 1987 to 77 percent in 1988, and to 84 percent in 1989. The mathematics test again was the most difficult for New Jersey students.

Given the recent trends regarding the overall success rate for ninth grade first-takers together with pass rates for reexaminees in subsequent years (ranging from 49 percent to 67 percent), one would expect almost all New Jersey students to satisfy the standards by graduation. Mathematics would seem to be the primary stumbling block for a small number of students and one would hypothesize that those few with poor ninth grade performance would be among those most likely to drop out of school.

In summary, ninth grade New Jersey students showed considerable growth in basic skills achievement on the HSPT examination from 1984 through 1989. Students correctly answered a very high percentage of items on all subtests. These data suggest that most New Jersey students will probably fulfill the HSPT graduation requirement by the completion of twelfth grade.

Pennsylvania

Pennsylvania first administered the Testing for Essential Learning and Literacy Skills (TELLS) examination package in the 1984-85 school year. This

package essentially is a minimum basic skills test in reading and mathematics administered to grades 3, 5, and 8. The purpose of TELLS is to diagnose basic skill deficiencies at an early stage so that appropriate remediation can be provided. In constructing the examinations, the state selected commercially available test items that matched their specified basic skills objectives. Thus, national normative data are available. Expert panels determined the passing scores for the tests in 1985 and 1986; the 1986 standard was used in subsequent years. Results of the TELLS examinations since 1985 are summarized in Table 7.

At the elementary level, the percentages of third grade students below the reading and mathematics cutoff scores have declined since the baseline year, indicating improved student performance. Reading performance has improved slowly and steadily each year, in terms of the percentage of students below the cutoff. At grade 5, the percentage of students below the cutoff has fluctuated somewhat over the years, with relatively high percentages in 1987 and 1989. Student performance in mathematics from 1985 to 1989 improved, although 1989 performance was much lower than 1988.

At the intermediate level, the percentages below the cutoff for reading have fluctuated for grade 8 students, with an overall improvement from 1985 to 1989, but a decline since last year. The percentages below the cutoff for mathematics for grade 8 students has increased, with a marked decline since last year.

Changes in examination items each year may explain the differences in the percentage of students above and below the cutoff score. While the cutoff scores remained the same from 1986 through 1989, the average difficulty level of the examinations varied in some cases. For example, national norms suggest

Table 7

Pennsylvania Achievement Score Trends
TELLS
(Percent Correct Scores)

Subtest	1985			1986			1987			1988			1989		
	% Below Cutoff Score	PA Aver. Score	Natl. Aver.	% Below Cutoff Score	PA Aver. Score	Natl. Aver.	% Below Cutoff Score	PA Aver. Score	Natl. Aver.	% Below Cutoff Score	PA Aver. Score	Natl. Aver.	% Below Cutoff Score	PA Aver. Score	Natl. Aver.
Grade 3															
Reading	26.8	74	72	24.2	79	78	23.2	79	78	23.6	79	78	21.8	78	78
Mathematics	19.2	78	78	18.8	83	81	16.6	85	81	14.0	86	83	15.2	84	79
Grade 5															
Reading	20.1	78	72	21.4	80	79	26.0	78	76	23.6	80	77	26.0	76	76
Mathematics	26.5	74	77	19.2	80	79	16.0	82	80	16.8	82	80	22.0	79	78
Grade 8															
Reading	24.7	76	76	21.7	79	78	23.7	78	77	19.2	80	81	20.6	79	81
Mathematics	22.2	70	68	24.8	78	76	25.1	77	76	21.7	79	77	23.5	77	74

differences in test difficulty from year to year (e.g., 1987 mathematics tests), thus helping to explain some apparent decline in student performance.

Pennsylvania students scored higher than national norms in mathematics each year since 1986, particularly in grades 3 and 8 for 1989. Average reading scores were at the national average for grades 3 and 5, but below the national norm in grade 8. Reading performance relative to national norms seemed to decline for all grade levels in 1989 compared to last year's performance.

In summary, TELLS performance of Pennsylvania students in 1989 was not as strong as 1988 performance, in general. For each grade and subtest, there was either a decline from last year relative to national norms or in the percentage of students above the cutoff. However, 1988 performance was very good and, for the most part, 1989 results are superior to 1985 results. Student performance in 1989 relative to national norms is very good in mathematics, but 1989 performance in reading is either at or below the national norm.

Analysis of Regional Basic Skills Achievement

The preceding section presented a state-by-state summary of achievement performance trends. This section briefly summarizes basic skills achievement across the region. These data allow only gross comparisons of student performance trends across the region. As previously noted, these results can be used only for within state and trend comparisons across the region and not as a means of determining whether one jurisdiction is performing better or worse than another. Trends are particularly difficult to discern in 1989 due to changes in several testing programs. Results are presented separately below for reading and mathematics. To provide some additional perspective for regional findings, recent results of the Scholastic Aptitude Test (SAT) are also presented.

Regional Reading Trends

Reading achievement trends across the region are summarized in Table 8 graphically and Table 9 narratively. Overall, there was a slight positive trend in reading achievement at the elementary grade level across the six-year period. There was also some evidence of slight declines in achievement at various points. However, scores generally improved from 1986-87, remained stable in 1988, and improved slightly in 1989. Grade 3 performance tended to be superior to grade 5 performance over the six-year period. Achievement was at the national norms in two states (PA and DC), slightly above the national norm in one state (DE), and considerably above in one state (MD). However, all norming samples were different, probably accounting for the variability in performance across the region. There was a high degree of achievement with

Table 8

Regional Reading Achievement Trends
Gross Trend Indicators (Trends Relative to Prior Year's Performance)

Elementary

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	3	*	-	++	-	+	+
DE	3	*	+	0	+	-	N.A.
MD	3	*	0	0	+	0	+
PA-EQA	4	N.A.	N.A.	*	++	0	N.A.
PA-TELLS	3	N.A.	*	++	+	0	+

Intermediate

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	8	*	0	++	--	++	0
DE	8	*	++	--	++	+	N.A.
MD	8	*	-	0	+	0	0
PA-EQA	7	N.A.	N.A.	*	+	+	N.A.
PA-TELLS	8	N.A.	*	++	-	++	-

Secondary

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	11	*	++	++	+	-	++
DE	11	*	++	-	0	+	N.A.
NJ	9	*	++	++	++	++	++
PA-EQA	11	N.A.	N.A.	*	+	-	N.A.

Note: Key to Gross Trend Indicators:

0:	No change from prior year
+ or - :	Change of 1-2 standard score points, or 1-2 percentage points (TELLS) from prior year
++ or --:	Change of 3 or more standard score points, or percentage points from prior year.
*:	Indicates baseline year.
N.A.:	Not available.

These arbitrary indicators are intended to give a general view of whether annual average test scores seem to be stable, improving, or declining relative to the previous year's scores; they should not be regarded as precise measures of performance.

Table 9

Summary of Reading Achievement Trends Across Region

Jurisdiction	Elementary	Intermediate	Secondary
District of Columbia (DC)	- Stable scores 84-88 - Slight improvement in both 1988 and 1989 - Scores at national norm	- Slight improvement 1987-88 - 1989 scores identical to 1986 and 1988 high - Scores lower than national norms 84-89	- Very low 1984 baseline score - Substantial improvement 1985-87, 1988-89 - 1988 scores stable with 1987 level - Scores much lower than national norms - SAT Verbal scores below national average; slight decline since 1986 after substantial improvement 84-86; slight improvement 1988-89
Delaware (DE)	- Overall slight improvement 84-87, with a slight decline 1987-88 - Moderately above national norms in 1989	- Overall, slight improvement 1984-88 - Slight decline 1985-86, improvement 1986-87 - Moderately above national norms in 1989	- Overall, slight improvement 1984-86 - Decline 1985-86, return to 1985 high in 1988 - Moderately above national norms in 1989 - SAT Verbal scores above national average; substantial decline in 1988, after improvement in 1985; slight improvement in 1988-89
Maryland (MD)	- High baseline scores - Slight decline or stable scores 1984-86 - Considerable improvement 1986-1989 - Grade 3 scores higher 1987-89 than in 1984-86 - Grade 5 scores stable 1985-89 after decline from 1984-85 - Scores much higher than national norms	- High baseline scores - Slight decline 1984-86 - Improvement 86-89, back to 84 level - Scores well above national norms	- No test data available - SAT Verbal scores above national average; steady improvement 1984-87, decline 1988; slight improvement in 1989

Jurisdiction	Elementary	Intermediate	Secondary
New Jersey (NJ)	- No test data available	- No test data available	- Considerable improvement in proficiency scores 1984-89 - Almost all students meet graduation requirements - SAT Verbal scores slightly below national average; relatively stable since 1985
Pennsylvania (PA)	- Slight improvement 1986-88 (EQA) - Improvement Grade 3 TELLS in 1989, decline for Grade 5 - Minimum skills at national norm (TELLS) - Approximately 76% of students meet minimum standards 1989 (TELLS)	- Slight improvement 1985-88 (EQA) - Improvement 1985-89 TELLS, despite decline in 1989 - Below national norm (TELLS) - Approximately 79% of students meet minimum standards 1989 (TELLS)	- Stable scores 1986-87 (EQA) - SAT Verbal scores slightly below national average; slight decline in 1988 and 1989 after improvement in 1985-1986

regard to minimum basic skills (NJ and PA). Most students met minimum proficiency standards set by these states.

Trends at the intermediate grade level were generally positive, with moderate overall improvement across the six-year period. Scores generally improved from 1984 through 1988 and were stable or declined slightly in 1989. Achievement was slightly below the national norm in two states (DC and PA), slightly above in one state (DE), and well above the norm in one state (MD). Most students met minimum proficiency standards.

Secondary level reading trends are somewhat difficult to assess since six year data were available for only two states. Performance in New Jersey has improved markedly each year and District of Columbia performance has, likewise, improved considerably each year except one. Pennsylvania and Delaware scores were relatively stable through 1988.

Scores on the Scholastic Aptitude Test are summarized in Table 10. Verbal scores on the SAT tended to increase from 1984 to 1985, peak in 1985 or 1986, and decline slightly in 1988. In general, 1989 scores were relatively stable, with a slight increase for three states and a slight decline for two. Overall, verbal scores within the region are below the national average (presented at the bottom of Table 10) in 1989, with the exceptions of Delaware and Maryland which have consistently been above the national average.

Table 10

Statewide Averages on Scholastic Aptitude Test

Verbal

Jurisdiction	1984	1985	1986	1987	1988	1989
DC	397	410	413	407	405	407
DE	433	444	442	440	433	435
MD	429	435	436	437	433	434
NJ	418	425	424	425	424	423
PA	425	428	429	428	424	423
National Average	426	431	431	430	428	427

Quantitative

Jurisdiction	1984	1985	1986	1987	1988	1989
DC	426	434	439	435	434	439
DE	469	474	475	470	466	468
MD	468	475	475	477	475	480
NJ	458	464	465	467	469	471
PA	462	465	465	463	462	463
National Average	471	475	475	476	476	476

In summary, there was some fluctuation in regional reading achievement trends from 1984 to 1989. Overall results were mixed with some jurisdictions making small gains and others occasional declines. However, average scores tended to be above national norms, with the exception of the District of Columbia and Grade 8 in Pennsylvania, where they were below the national average.

Regional Mathematics Trends

Mathematics achievement trends are summarized in Table 11 graphically and Table 12 narratively. There was general improvement in mathematics achievement across the region from 1984 to 1989. Gains were substantial in some cases. Trends were similar for all grade levels. Much of the improvement during the period occurred from 1986 through 1988. However, changes from 1988 to 1989 varied across states. While there was some evidence of improvement, there were also some declines. Maryland elementary level scores declined in 1989 as did Pennsylvania elementary and intermediate scores.

Scores, in general, were considerably higher than national norms, with the exception of the District of Columbia intermediate and secondary level scores. Also, there was a high degree of achievement with regard to minimum basic skills in mathematics. For example, most students in Pennsylvania and New Jersey met minimum proficiency standards established by these states.

Quantitative scores on the Scholastic Aptitude Test have remained moderately (DE, NJ, PA) to well below (DC) the national average over the five year period. Only Maryland has consistently scored at or close to the national average. However, scores in all states improved in 1989. Trends in the District of Columbia, Maryland, and New Jersey have improved since 1984, while Pennsylvania and Delaware scores have remained stable.

Table 11

Regional Mathematics Achievement Trends
Gross Trend Indicators (Trends Relative to Prior Year's Performance)

Elementary

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	3	*	-	+	0	++	++
DE	3	*	0	++	+	--	N.A.
MD	3	*	0	0	++	+	-
PA-EQA	4	N.A.	N.A.	*	++	++	N.A.
PA-TELLS	3	N.A.	*	0	+	++	-

Intermediate

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	8	*	++	+	-	+	+
DE	8	*	++	-	+	+	N.A.
MD	8	*	-	+	+	+	0
PA-EQA	7	N.A.	N.A.	*	++	++	N.A.
PA-TELLS	8	N.A.	*	--	0	++	-

Secondary

Jurisdiction	Grade	1984	1985	1986	1987	1988	1989
DC	11	*	+	+	++	0	+
DE	11	*	+	0	+	+	N.A.
NJ	9	*	+	++	++	+	++
PA-EQA	11	N.A.	N.A.	*	0	+	N.A.

Note: Key to Gross Trend Indicators:

0:	No change from prior year
+ or - :	Change of 1-2 standard score points, or 1-2 percentage points (TELLS) from prior year
++ or --:	Change of 3 or more standard score points, or percentage points from prior year.
*:	Indicates baseline year.
N.A.:	Not available.

These arbitrary indicators are intended to give a general view of whether annual average test scores seem to be stable, improving, or declining relative to the previous year's scores; they should not be regarded as precise measures of performance.

Table 12

Summary of Mathematics Achievement Trends Across Region

Jurisdiction	Elementary	Intermediate	Secondary
District of Columbia (DC)	- Declined slightly 1984-86, grade 3 and 6 - Improved 1986-89, to a level higher than baseline scores - Considerable improvement in 1989 - Above national norms in 1989	- Moderate improvement 1984-86 - Slight decline 1987, slight gains in 1988 and 1989 - Slightly below national norms	- Moderate improvement 1984-87 - Stable 1987-89 - Below national norms - SAT Quantitative scores well below national average; improvement 1984-86, decline 1986-88, improvement 1988-89
Delaware (DE)	- Steady improvement 1984-87, slight decline in 1988 - Scores slightly above national norms in 1989	- Moderate improvement 1984-88 - Slightly above national norms in 1989	- Moderate improvement 1984-88 - Slightly above national norms in 1989 - SAT Quantitative scores below national average; improvement 1984-86, decline 1986-88, slight improvement in 1989
Maryland (MD)	- Stable 1984-86 - Improvement 1986-87 - Stable 1987-89 - Well above national norms	- Slight improvement 1984-88 - Stable 1989 - Well above national norms	- No test data available - SAT Quantitative scores above national norm in 1989; improvement 1984-89, particularly in 1989
New Jersey (NJ)	- No test data available	- No test data available	- Considerable improvement in proficiency scores 1984-89, particularly latter years - Most students meet graduation proficiency standards - SAT Quantitative scores slightly below national average; steady improvement each year 1984-89

Jurisdiction	Elementary	Intermediate	Secondary
Pennsylvania (PA)	- Substantial improvement 1986-88 (EQA) - Improvement 1985-89 TELLS, despite decline in 1989 - Scores higher than national average (TELLS) - Approximately 82% of students meet minimum standards 1989 (TELLS)	- Substantial improvement grade 8, stable 1984-88 (EQA) - Scores higher than national norms (TELLS) - Approximately 76% of students meet minimum standards 1989 (TELLS) - Decline 1989 TELLS; 1989 performance lower than 1985 performance	- Improved scores 1986-88, grade 11 (EQA) - SAT Quantitative scores below national average; improvement 1984-85, relatively stable 1987-89

Conclusions

This report has summarized test results from the five jurisdictions in the RBS region. Performance trends over the past several years were analyzed. Major conclusions are briefly noted below.

- Reading achievement trends fluctuated somewhat across the region from 1984-1989. Overall performance trends were mixed. In general, 1989 scores improved from 1988 levels for elementary and secondary grades but remained stable or declined at the intermediate grade level.

- Students at all grade levels throughout the region showed overall improvement in mathematics during the overall period from 1984-1989. However, there was some evidence of declines from 1988 to 1989, particularly in Maryland and Pennsylvania.

- While 1989 test scores were generally positive, overall results were not as strong as in most prior years.

- Overall, students across the region perform higher than national norms on standardized tests. Scores were particularly high at the elementary school level. There were some exceptions for District of Columbia and Pennsylvania students at some grade levels.

- There was a high degree of achievement with regard to minimum basic skills in reading and mathematics. Most students met minimum proficiency standards established by the states.

- Overall, scores on both the verbal and quantitative subtests of the Scholastic Aptitude Test (SAT) tended to be below the national average, with the exception of Delaware (verbal) and Maryland (verbal and quantitative). This can be a deceiving statistic, however, since more students, representing a wide range of achievement trends, take the SATs than students from states in other parts of the country. In general, regional SAT scores improved from 1988 to 1989.

In summary, regional achievement trends were positive during the period from 1984 through 1989. These trends generally followed those exhibited from 1978-1984. Findings were similar to other reports of national trends (e.g., Wiser & Lenke, 1987). While the overall trends were positive, some annual fluctuations were noted. In some cases, apparent declines may possibly be attributed to change in test forms. These declines should be a concern if they are persistent; to date, they have not been so. Although the aggregate

score trends were positive, there was some variability within states. Improvement in achievement can probably be attributed to greater focus on school improvement and on the objectives of the test, greater familiarity on the part of teachers and students with the tests themselves, including the format and other administration factors. In the case of the testing programs that use national norms, some of the apparent improvement is due to the use of norming samples from many years ago, as noted in the Cannell reports. Still, districts throughout the region have focused considerable effort on school improvement in recent years and these have greater focus on the objectives measured by the tests. Given the recent positive performance trends and the increased focus on accountability, each of the states in the region has just completed or is in the process of carefully reviewing and revising their testing programs so that accurate indicators of achievement are available for students, parents, and educators.

References

Biester, T. W., & Buttram, J. L. (1986). Trends in school improvement: Regional test results 1978-1986. Philadelphia: Research for Better Schools, Inc.

Biester, T. W., & Dusewicz, R. A. (1982). Trends in school improvement: Statewide test results. Philadelphia: Research for Better Schools, Inc.

Biester, T. W., & Dusewicz, R. A. (1983). Trends in school improvement: Statewide test results, 1978-1983. Philadelphia: Research for Better Schools, Inc.

Biester, T. W., Herzel-Colvson, K., & Buttram, J. L. (1987). Trends in school improvement: Regional test results, 1984-1987. Philadelphia: Research for Better Schools, Inc.

Biester, T. W., & Rioux, S. C. (1988). Trends in school improvement: Regional test results, 1984-1988. Philadelphia: Research for Better Schools, Inc.

Cannell, J. J. (1980). The Lake Wobegan Effect Revisited. Educational Measurement: Issues and Practice, Winter, 12-15.

College Board (1988). Number of SAT candidates with verbal and math mean scores and standard deviations for each state from 1971-7 through 1987-88. Princeton, NJ: Author.

Kenney, J. L., & Dusewicz, R. A. (1984). Trends in school improvement: Statewide test results, 1978-1984. Philadelphia: Research for Better Schools, Inc.

Wiser, B. & Lenke, J. M. (1987). The stability of achievement norms over time. Paper presented at the Annual Meeting of the American Educational Research Association, April, 1987.